

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Tong Zhang  
Serial No. : 10/611,449  
Filed : July 1, 2003  
Title : AUDIO SUMMARY BASED AUDIO PROCESSING

Art Unit : 2615  
Examiner : Swerdlow, Daniel  
Confirmation No. 1631

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

APPEAL BRIEF

I. Real Party in Interest

The real party in interest is Hewlett-Packard Development Company, L.P., a Texas Limited Partnership having its principal place of business in Houston, Texas.

II. Related Appeals and Interferences

Appellant is not aware of any related appeals or interferences that will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

III. Status of Claims

Claims 1-17 and 39-62 are pending.

Claims 18-38 have been canceled.

Of the pending claims, claims 1, 2, 5-17, and 39-62 are rejected, claim 3 is allowed, and claim 4 is objected to as being dependent upon a rejected base claim, but would be allowable if

CERTIFICATE OF TRANSMISSION

I hereby certify that this document is being transmitted to the Patent and Trademark Office via electronic filing.

Oct. 8, 2008

Date of Transmission

/Edouard Garcia, Reg. No. 38,461/

(Signature of person mailing papers)

Edouard Garcia

(Typed or printed name of person mailing papers)

rewritten in independent form including all of the limitations of the base claim and any intervening claims.

Appellant appeals all rejections of the rejected claims 1, 2, 5-17, and 39-62.

#### IV. Status of Amendments

The amendments filed February 14, 2008, have been entered and acted upon by the Examiner.

No amendments were filed after the final Office action dated July 8, 2008.

#### V. Summary of Claimed Subject Matter

In the following Summary, the citations in parentheses are representative of support provided in the application.

##### A. Independent claim 1

The aspect of the invention defined in independent claim 1 is an audio processing method in accordance with which audio summaries of respective audio pieces are identified (page 6, line 15 - page 7, line 4; FIG. 2). Each of the audio summaries comprises digital content summarizing at least a portion of the respective audio piece (page 4, lines 14-21). The process of identifying the audio summaries comprises for each of the audio pieces: selecting constituent segments of the audio piece as its respective ones of the audio summaries (page 6, line 29 - page 7, line 2; FIG. 2, block 68); and ranking its audio summaries into different levels of a respective audio summary hierarchy (page 7, lines 2-4; page 10, lines 24-31; page 12, lines 16-22). Transition audio segments are determined, where each transition audio segment comprises a form of audio content that is different from the audio summaries and distinguishes the transition audio segment from the audio summaries (page 11, lines 5-7 and 22-24). The transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies are concatenated into a sequence in which at least one of the transition audio segments is between

successive ones of the audio summaries (page 11, lines 5-12; Fig. 7). The sequence is rendered (page 12, lines 5-15).

**B. Dependent claim 10**

Claim 10 depends from claim 9 and recites rendering a given one of the audio pieces linked by a browsable link to an associated one of the audio summaries in response to user input received during rendering of the associated audio summary (page 3, lines 7-11; page 7, lines 16-26; page 8, lines 27-31), wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece before rendering a successive one of the transition audio segments (page 9, line 10 - page 10, line 10; FIG. 5; page 12, line 31 - page 13, line 17; FIG. 8).

**C. Dependent claim 11**

Claim 11 depends from claim 1 and recites rendering a given audio piece beginning at a location in the given audio piece linked by a browsable link to an audio summary associated with the given audio piece, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece (page 3, lines 7-11; page 7, lines 16-26; page 8, lines 27-31; page 9, line 10 - page 10, line 10; FIG. 5; page 12, line 31 - page 13, line 17; FIG. 8).

**D. Dependent claim 16**

Claim 16 depends from claim 1 and recites normalizing audio summaries to a common loudness level (page 11, lines 12-14).

**E. Dependent claim 46**

Claim 46 depends from claim 1 and recites following a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the specified location (page 10, line 32 - page 11, line 3; FIG. 6, block 139; page 12, line 31 - page 13, line 17; FIG. 8).

F. Dependent claim 47

Claim 47 depends from claim 46 and recites terminating the rendering of the associated audio piece and resuming the sequential rendering of the audio summaries and the transition audio segments (page 12, lines 16-22).

G. Dependent claim 49

Claim 49 depends from claim 47 and recites that the terminating is initiated in response to completion of the rendering of the associated audio piece (page 13, lines 10-17; FIG. 8).

H. Independent claim 13

The aspect of the invention defined in independent claim 13 is an audio processing method in accordance with which audio summaries and transition audio segments are sequentially rendered with at least one transition audio segment rendered between each pair of sequential audio summaries (page 11, lines 5-21). Each of the audio summaries comprises digital content summarizing at least a portion of a respective associated audio piece (page 4, lines 14-21). A user request to browse the audio summaries is received (page 10, lines 22-31; FIG. 6, block 138). Ones of the audio summaries are ordered into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was received (page 7, lines 2-4; page 10, lines 24-31; page 12, lines 16-22; page 13, line 28 - page 14, line 4). The sequence is rendered (page 12, lines 16-22).

I. Independent claim 15

The aspect of the invention defined in independent claim 15 is an audio processing method in accordance with which audio summaries and transition audio segments are sequentially rendered with at least one transition audio segment rendered between each pair of sequential audio summaries (page 11, lines 5-21). Each audio summary comprises digital content summarizing at least a portion of a respective associated audio piece (page 4, lines 14-21). Each audio piece is associated with multiple audio summaries and a single audio summary is rendered automatically for each audio piece (page 11, lines 15-21; FIG. 7). An audio

summary is rendered for a given audio piece in response to user input received during rendering of a preceding audio summary associated with the given audio piece (page 12, lines 16-22).

J. Independent claim 17

The aspect of the invention defined in independent claim 17 is an audio processing system, comprising a rendering engine (FIG. 1, block 12) operable to perform operations comprising the following operations. Audio summaries of respective audio pieces are identified (page 6, line 15 - page 7, line 4; FIG. 2). Each of the audio pieces is associated with respective ones of audio summaries ranked into different levels of a respective audio summary hierarchy and in the identifying the rendering engine is operable to identify the respective levels into which the audio summaries are ranked (page 7, lines 2-4; page 10, lines 24-31; page 12, lines 16-22). Transition audio segments are determined, wherein each transition audio segment comprises a form of audio content that is different from the audio summaries and distinguishes the transition audio segment from the audio summaries (page 11, lines 5-7 and 22-24). The transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies are concatenated into a sequence in which at least one of the transition audio segments is between sequential ones of the audio summaries (page 11, lines 5-12; Fig. 7). The sequence is rendered (page 12, lines 5-15).

K. Independent claim 52

The aspect of the invention defined in independent claim 52 is an audio processing system that comprises a rendering engine (FIG. 1, block 12). The rendering engine is operable to sequentially render audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries (page 11, lines 5-21), wherein each of the audio summaries comprises digital content summarizing at least a portion of a respective associated audio piece (page 4, lines 14-21). In response to receipt of a user request to browse the audio summaries (page 10, lines 22-31; FIG. 6, block 138), the rendering engine is operable to order ones of the audio summaries into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was

received and to render the sequence (page 7, lines 2-4; page 10, lines 24-31; page 12, lines 16-22; page 13, line 28 - page 14, line 4).

L. Independent claim 54

The aspect of the invention defined in independent claim 54 is an audio processing system that comprises a rendering engine operable to sequentially render audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries (page 11, lines 5-21), wherein each audio piece is associated with multiple audio summaries (page 11, lines 15-21; FIG. 7). The rendering engine is operable to render a single audio summary automatically for each audio piece (page 11, lines 15-21; FIG. 7). The rendering engine additionally is operable to render an audio summary for a given audio piece in response to user input received during rendering of a preceding audio summary associated with the given audio piece (page 12, lines 16-22).

VI. Grounds of Rejection to be Reviewed on Appeal

A. Claims 1, 2, 5, 6, 9-12, 17, 39-41, 45-48, 55, and 58-62 stand rejected under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047).

B. Claims 7, 8, 15, 42-44, 50, 51, 54, 56, and 57 stand rejected under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Csicsatka (U.S. 2006/0235550).

C. Claims 13, 14, 52, and 53 stand rejected under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Herley (U.S. 2004/0002310).

D. Claim 16 stands rejected under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Takenaka (U.S. 6,807,450).

E. Claim 49 stands rejected under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Setogawa (U.S. 6,424,793).

## VII. Argument

### **A. Applicable standards for sustaining a rejection under 35 U.S.C. § 103(a)**

“A patent may not be obtained ... if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.” 35 U.S.C. §103(a).

In an appeal involving a rejection under 35 U.S.C. § 103, an examiner bears the initial burden of establishing *prima facie* obviousness. See In re Rijckaert, 9 F.3d 1531, 1532, 28 USPQ2d 1955, 1956 (Fed. Cir. 1993). To support a *prima facie* conclusion of obviousness, the prior art must disclose or suggest all the limitations of the claimed invention.<sup>1</sup> See In re Lowry, 32 F.3d 1579, 1582, 32 USPQ2d 1 031, 1034 (Fed. Cir. 1994). If the examiner has established a *prima facie* case of obviousness, the burden of going forward then shifts to the applicant to overcome the *prima facie* case with argument and/or evidence. Obviousness, is then determined on the basis of the evidence as a whole and the relative persuasiveness of the arguments. This inquiry requires (a) determining the scope and contents of the prior art; (b) ascertaining the differences between the prior art and the claims in issue; (c) resolving the level of ordinary skill in the pertinent art; and (d) evaluating evidence of secondary consideration. See KSR Int'l Co. v. Teleflex Inc., No. 127 S. Ct. 1727, 1728 (2007) (citing Graham v. John Deere, 383 U.S. 1, 17-18, 148 USPQ 459, 467 (1966)). If all claim limitations are found in a number of prior art references, the fact finder must determine whether there was an apparent reason to combine the known elements in the fashion claimed. See KSR, 1741. This analysis should be made explicit.

---

<sup>1</sup> The U.S. Patent and Trademark Office has set forth the following definition of the requirements for establishing a *prima facie* case of unpatentability (37 CFR § 1.56(b)(ii):

A *prima facie* case of unpatentability is established when the information compels a conclusion that a claim is unpatentable under the preponderance of evidence, burden-of-proof standard, giving each term in the claim its broadest reasonable construction consistent with the specification, and before any consideration is given to evidence which may be submitted in an attempt to establish a contrary conclusion of patentability.

KSR at 1741 (citing In re Kahn, 441 F. 3d 977, 988 (Fed. Cir. 2006): “[R]ejections on obviousness grounds cannot be sustained by mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness”).

**B. The rejection of claims 1, 2, 5, 6, 9-12, 17, 39-41, 45-48, 55, and 58-62 under 35 U.S.C. § 103(a) over Jun in view of Kulas**

The Examiner has rejected claims 1, 2, 5, 6, 9-12, 17, 39-41, 45-48, 55, and 58-62 under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047).

**I. Independent Claim 1**

**a. Introduction**

Independent claim 1 recites:

1. An audio processing method, comprising:
  - identifying audio summaries of respective audio pieces, wherein each of the audio summaries comprises digital content summarizing at least a portion of the respective audio piece, and the identifying comprises for each of the audio pieces
    - selecting constituent segments of the audio piece as its respective ones of the audio summaries and
    - ranking its audio summaries into different levels of a respective audio summary hierarchy;
    - determining transition audio segments each comprising a form of audio content that is different from the audio summaries and distinguishes the transition audio segment from the audio summaries;
    - concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence in which at least one of the transition audio segments is between successive ones of the audio summaries; and
    - rendering the sequence.



The rejection of independent claim 1 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn because Jun and Kulas, taken either alone or in any permissible combination, do not disclose or suggest all the elements of the claimed invention. The rejection of independent claim 1 also should be withdrawn because at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas in the manner proposed by the Examiner.

b. The Examiner's position

In support of the rejection of claim 1, the Examiner has taken the following position (see § 3 on pages 2-3 of the final Office action):

- Jun discloses each and every one of the elements of claim 1 except “is silent as to any transitional audio segments and browsing implementations”;
- Kulas makes-up for Jun's “silence” in col. 5, lines 52-65; and
- “Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to apply the transition audio segments taught by Kulas in the digest streams taught by Jun...”

c. Appellant's rebuttal: the cited references do not disclose each and every element of the invention defined in claim 1

Contrary to the Examiner's position, Jun and Kulas, taken either alone or in any permissible combination, do not disclose or suggest all the elements of the claimed invention.

For example, neither Jun nor Kulas discloses or suggests “concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence in which at least one of the transition audio segments is between successive ones of the audio summaries,” where the audio summaries of each of the audio pieces are ranked into a respective one of the audio summary hierarchies, as recited in claim 1.

Jun discloses a method in which digest segments of a multimedia stream of a particular multimedia content (e.g., movie, drama, sports; see col. 1, line 29) are identified and ranked into different levels of a digest segment hierarchy (see col. 4, lines 20-29, and FIG. 1), where each digest segment is a sub-stream of the original multimedia stream (see col. 1, lines 35-36 and 56-

57). Regarding the way in which the digest segments extracted from a single multimedia content are processed, Jun discloses that in response to a user-selected running time the determining unit 1 (FIG. 6) determines the digest level of the digest segment hierarchy for the particular multimedia content and outputs the digest segments at the determined digest level to the decoder 3, which outputs the decoded segments to the browsing/editing/recording unit 4 (see col. 6, lines 43-53). Jun discloses that "The outputted digest segments are sequentially browsed, recorded, and edited by the browsing/recording/editing unit 4" (col. 6, lines 51-53). Although Jun is silent on the issue, one reasonably can infer from the fact that the particular multimedia content is received as a stream (see, e.g., FIG. 6) that the digest segments of the particular multimedia content are sequentially output from the determining unit 1, sequentially decoded by the decoder 3, and the decoded segments are sequentially recorded in a concatenated sequence in the browsing/recording/editing unit 4. Jun does not disclose anything whatsoever about concatenating the digest segments at a selected level of multiple digest segment hierarchies extracted from respective multimedia contents.

Thus, as acknowledged by the Examiner Jun does not disclose or suggest "concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence in which at least one of the transition audio segments is between successive ones of the audio summaries," as recited in claim 1. Contrary to the Examiner's position, however, Kulas does not make-up for the failure of Jun to disclose or suggest "concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence in which at least one of the transition audio segments is between successive ones of the audio summaries," as recited in claim 1.

In pertinent part Kulas discloses a multi-CD player 100 that includes a controller 118 that stores in RAM a four-second sample of the beginning of each CD that is loaded in the player 100 and a tone generator 154 that outputs a tone "to indicate the end of a sample during sample playback mode and to indicate the start of a next sample" (col. 5, lines 53-56). Kulas discloses selecting a single four-second sample as the summary of all of audio pieces on a respective CD, and rendering a single four-second sample of the beginning of each CD in order of CD slot number (see, e.g., col. 5, lines 21-49; FIG. 3).

Kulas does not disclose or suggest anything about selecting multiple audio summaries for each of the audio pieces, nor does Kulas disclose or suggest anything about ranking multiple audio summaries of an audio piece into different levels of a respective audio summary hierarchy. Consequently, Kulas cannot possibly disclose anything whatsoever about concatenating the digest segments at a selected level of multiple digest segment hierarchies extracted from respective multimedia contents.

Thus, neither Jun nor Kulas discloses the "concatenating" element of claim 1. Since the cited references do not disclose or suggest each and every one of the elements of claim 1, the rejection of claim 1 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn.

It is noted that the combination of Jun and Kulas proposed by the Examiner does not contain each and every element of claim 1. In particular, The Examiner has taken the position that "...it would have been obvious to one of ordinary skill in the art at the time of the invention to apply the transition audio segments taught by Kulas in the digest streams taught by Jun..." As explained above, Jun discloses a method in which digest segments of a multimedia stream of a single multimedia content are identified and ranked into different levels of a digest segment hierarchy, and Kulas discloses outputting a tone "to indicate the end of a sample during sample playback mode and to indicate the start of a next sample" (col. 5, lines 53-56). Thus, the Examiner's proposed combination of Jun and Kulas would result in an audio tone being played at the end of each digest segment at a selected level of a single digest segment hierarchy extracted from a single multimedia content. The proposed combination, however, would not result in concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies (plural) into a sequence in which at least one of the transition audio segments is between successive ones of the audio summaries, as recited in claim 1.

d. Appellant's rebuttal: at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas in the manner proposed by the Examiner

Contrary to the Examiner's position, at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas in the manner proposed by the Examiner.

The Examiner has taken the position that (see § 2 on page 4 of the final Office action):

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to apply the transition audio segments taught by Kulas in the digest streams taught by Jun, since the transitional audio segments are used "to indicate the end of a sample during playback mode and to indicate the start of a next sample," Kulas Column 5 Lines 52 -65, thereby allowing the user do distinguish when one sample ends and the other begins. In addition to the transitional segments allowing for smooth browsing as taught by Kulas, it would have also been obvious to one of ordinary skill in the art at the time of the invention to incorporate other aspects relating to browsing as taught by Kulas in the invention of Jun, since Jun also states "Meanwhile, the operation of browsing, recording and editing performed by the browsing/recording/editing unit 4 is carried out by a well-known technique of the convention art, and accordingly a detailed description thereof is omitted," Column 6 Lines 20 -26.

As explained above, Jun discloses a method in which digest segments of a multimedia stream of a particular multimedia content (e.g., movie, drama, sports; see col. 1, line 29) are identified and ranked into different levels of a digest segment hierarchy (see col. 4, lines 20-29, and FIG. 1), where each digest segment is a sub-stream of the original multimedia stream (see col. 1, lines 35-36 and 56-57). Kulas discloses outputting a tone "to indicate the end of a sample during sample playback mode and to indicate the start of a next sample" (col. 5, lines 53-56). Thus, the Examiner's proposed combination of Jun and Kulas would result in a audio tone being played at the end of each digest segment at a selected level of a single digest segment hierarchy extracted from a single multimedia content.

For the reasons explained below, one skilled in the art the time the invention was made would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner.

First, one skilled in the art would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner because the proposed combination would not have served any readily apparent purpose. Jun discloses that digest segments of an audio-visual multimedia stream (e.g., movie, drama, sports; see col. 1, line 29) are identified and ranked into different levels of a digest segment hierarchy (see col. 4, lines 20-29, and FIG. 1), where each

digest segment is a sub-stream of the original multimedia stream (see col. 1, lines 35-36 and 56-57). The audio-video multimedia streams disclosed in Jun are quite difference from the audio pieces disclosed in Kulas. In particular, in the case of audio pieces, it is not readily readily apparent to the listener when a four-second audio sample of an audio piece has ended and a four-second audio sample of a successive audio piece has begun because the audio samples of different audio pieces from the same CD oftentimes sound very similar to one another. This is the reason Kulas inserts a tone between the audio samples. In the case of audio-video content, on the other hand, the video content itself provides all the cues needed to inform the viewer when a digest segment has ended because the video contents of successive digest segments selected from the same multimedia stream are quite likely to be visually different. This is the reason Jun does not disclose or suggest anything whatsoever about inserting audible tones between successive digest segments.

Second, one skilled in the art would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner because the proposed combination would produce an undesirable result. In particular, playing a tone as taught by Kulas between the audio-visual digest segments of a particular multimedia stream as taught by Jun would be quite discordant to the viewer who is browsing through sequences of the digest segments because such audio interruptions are not necessary and therefore are not expected in the context of audio-video browsing.

For at least this additional reason, the rejection of independent claim 1 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn.

d. Conclusion

For the reasons explained above, the rejection of independent claim 1 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn because Jun and Kulas, taken either alone or in any permissible combination, do not disclose or suggest all the elements of the claimed invention. The rejection of independent claim 1 also should be withdrawn because at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas in the manner proposed by the Examiner.

2. Claims 2, 5, 6, 9-12, 39-41, 45-48, 61, and 62

a. Introduction

Each of claims 2, 5, 6, 9-12, 39-41, 45-48, 61, and 62 incorporates the elements of independent claim 1 and therefore is patentable over Jun in view of Kulas for at least the same reasons explained above.

Claims 10, 11, 46, and 47 also are patentable over Jun in view of Kulas for the following additional reasons.

b. Claim 10

Claim 10 depends from claim 9 and recites that the method comprises rendering a given one of the audio pieces linked by a browsable link to an associated one of the audio summaries in response to user input received during rendering of the associated audio summary, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece before rendering a successive one of the transition audio segments.

In support of the rejection of claim 10, the Examiner has taken the position that (pages 5-6 of the final Office action):

Claim 10: Jun and Kulas disclose the method of claim 9, further comprising rendering a given one of the audio pieces linked by a browsable link to an associated one of the audio summaries in response to user input received during rendering of the associated audio summary, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece before rendering a successive one of the transition audio segments (Kulas Column 6 Lines 12 -27).

Contrary to the Examiner's position, however, col. 6, lines 12-27 of Kulas does not disclose or suggest rendering a given one of the audio pieces linked by a browsable link to an associated one of the audio summaries in response to user input received during rendering of the associated audio summary, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece before rendering a successive one of the transition audio segments.

In col. 6, lines 12-27, Kulas discloses that when the user presses the PLAY button 162 while a sample is being played, the system notes the slot corresponding to the current sample, retrieves the CD for the noted slot, and plays the retrieved CD. This process does not involve rendering a given one of the audio pieces linked by a browsable link to an associated one of the audio summaries in response to user input received during rendering of the associated audio summary, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece before rendering a successive one of the transition audio segments. Indeed, the CD slot is not a browsable link to an associated one of the audio samples; instead, the slot is the physical location where the CD is stored, and the samples are stored in RAM (see col. 3, lines 5-18), not the CD. In addition, the playing of the CD does not constitute following the browsable link from the associated audio summary to the given audio piece; instead, Kulas' multi-CD player plays the CD in the noted slot without reference to any browsable link from an audio sample to an audio piece on the CD.

For at least this additional reason, the rejection of claim 46 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn.

c. Claim 11

Claim 11 depends from claim 1 and recites that the method comprises rendering a given audio piece beginning at a location in the given audio piece linked by a browsable link to an audio summary associated with the given audio piece, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece.

In support of the rejection of claim 11, the Examiner has taken the position that (pages 5-6 of the final Office action):

Claim 11: Jun and Kulas disclose the method of claim 1, further comprising rendering a given audio piece beginning at a location in the given audio piece linked by a browsable link to an audio summary associated with the given audio piece, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece (Kulas Column 6 Lines 12 -27).

Contrary to the Examiner's position, however, col. 6, lines 12-27 of Kulas does not disclose or suggest rendering a given audio piece beginning at a location in the given audio piece linked by a browsable link to an audio summary associated with the given audio piece, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece.

In col. 6, lines 12-27, Kulas discloses that when the user presses the PLAY button 162 while a sample is being played, the system notes the slot corresponding to the current sample, retrieves the CD for the noted slot, and plays the retrieved CD. This process does not involve rendering a given audio piece beginning at a location in the given audio piece linked by a browsable link to an audio summary associated with the given audio piece, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece. Indeed, the CD slot is not a browsable link to an associated one of the audio samples; instead, the slot is the physical location where the CD is stored, and the samples are stored in RAM (see col. 3, lines 5-18), not the CD. In addition, the playing of the CD does not constitute following the browsable link from the associated audio summary to the given audio piece; instead, Kulas' multi-CD player plays the CD in the noted slot without reference to any browsable link from an audio sample to an audio piece on the CD.

For at least this additional reason, the rejection of claim 46 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn.

d. Claim 46

Claim 46 depends from claim 1 and recites following a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the specified location.

In support of the rejection of claim 46, the Examiner has stated that (see page 8 of the final Office action):

Claim 46: Jun and Kulas disclose the method of claim 1, further comprising following a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the



specified location (Kulas Column 3 Lines 27 -40 and Column 6 Lines 12 -27).

Contrary to the Examiner's statement, however, Kulas does not disclose or suggest following a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the specified location.

In col. 3, lines 27-40, Kulas discloses that a respective sample flag is used to indicate whether a valid sample exists for each slot in a multi-CD player; the flags are reset when the system is reset or when the corresponding CDs are swapped out, and the flags are set whenever a new sampling of a new CD completes. This disclosure has nothing whatsoever to do with following a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the specified location.

In col. 6, lines 12-27, Kulas discloses that when the user presses the PLAY button 162 while a sample is being played, the system notes the slot corresponding to the current sample, retrieves the CD for the noted slot, and plays the retrieved CD. This process does not involve following a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the specified location. Indeed, the CD slot does not point to a location in an associated audio piece, and the playing of the CD does not constitute rendering the associated audio piece beginning at a location in the audio piece specified by the CD slot; instead, Kulas' multi-CD player plays the CD in the noted slot without reference to any pointer from the CD slot to a location in an audio piece on the CD.

For at least this additional reason, the rejection of claim 46 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn.

e. Claim 47

Claim 47 depends from claim 46 and therefore is patentable over Jun in view of Kulas for the same additional reasons explained above in connection with claim 46. Claim 47 also is patentable over Jun in view of Kulas for the following additional reason.

Claim 47 recites terminating the rendering of the associated audio piece and resuming the sequential rendering of the audio summaries and the transition audio segments.

In support of the rejection of claim 47, the Examiner has stated that (see page 8 of the final Office action):

Claim 47: Jun and Kulas disclose the method of claim 46, further comprising terminating the rendering of the associated audio piece and resuming the sequential rendering of the audio summaries and the transition audio segments ("The SCAN control allows a user to initiate the quick-scan of the present invention," Kulas Column 5 Lines 7 -17, thereby "terminating" the current operating mode or the rendering of audio pieces).

Contrary to the Examiner's statement, however, Kulas does not disclose or suggest terminating the rendering of the associated audio piece and resuming the sequential rendering of the audio summaries and the transition audio segments. As noted by the Examiner, col. 5, lines 7-17 of Kulas discloses that the SCAN control allows a user to initiate the quick-scan feature. However, when the quick-scan feature is initiated it does not resume the sequential rendering of the audio summaries and the transition audio segments, as recited in claim 47. Instead, upon selection of the SCAN control, the system replays the entire set of CD samples from the beginning (see, e.g., col. 5, lines 18-51; FIG. 3).

For at least this additional reason, the rejection of claim 47 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn.

### 3. Independent claim 17

Independent claim 17 recites:

17. An audio processing system, comprising a rendering engine operable to perform operations comprising:

identifying audio summaries of respective audio pieces, wherein each of the audio pieces is associated with respective ones of audio summaries ranked into different levels of a respective audio summary hierarchy and in the identifying the rendering engine is operable to identify the respective levels into which the audio summaries are ranked;

determining transition audio segments each comprising a form of audio content that is different from the audio summaries and distinguishes the transition audio segment from the audio summaries;

concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence in which at least one of the transition audio segments is between sequential ones of the audio summaries; and

rendering the sequence.

Independent system claim 17 recites elements that essentially track the pertinent features of independent claim 1 discussed above. Therefore, claim 17 is patentable over Jun in view of Kulas for at least the same reasons explained above in connection with independent claim 1. In particular, the rejection of independent claim 1 under 35 U.S.C. § 103(a) over Jun in view of Kulas should be withdrawn because Jun and Kulas, taken either alone or in any permissible combination, do not disclose or suggest all the elements of the claimed invention. For example, neither Jun nor Kulas discloses or suggests the "concatenating" element of claim 17 for the reasons explained above in connection with independent claim 1. The rejection of independent claim 1 also should be withdrawn because at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas in the manner proposed by the Examiner, as explained above in connection with independent claim 1.

### 3. Claims 55 and 58-60

Each of claims 55 and 58-60 incorporates the elements of independent claim 17 and therefore is patentable over Jun in view of Kulas for at least the same reasons explained above.

#### **C. The rejection of claims 7, 8, 15, 42-44, 50, 51, 54, 56, and 57 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Csicsatka**

The Examiner has rejected claims 7, 8, 15, 42-44, 50, 51, 54, 56, and 57 under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Csicsatka (U.S. 2006/0235550, to the extent supported by the provisional patent application number 60/465,156).

1. Claims 7, 8, and 42-44

Each of claims 7, 8, and 42-44 incorporates the elements of independent claim 1.

Csicsatka does not make-up for the failure of Jun in view of Kulas to disclose and suggest the elements of independent claim 1 discussed above. Indeed, Csicsatka does not expressly or inherently disclose concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence. To the contrary, in accordance with Csicsatka's teachings the audio clips of an album are rendered without any type of transition between the audio clips.

Therefore, claims 7, 8, and 42-44 are patentable over Jun in view of Kulas and Csicsatka for at least the same reasons explained above in connection with independent claim 1.

2. Claims 50, 51, 54, 56, and 57

Each of claims 50, 51, 54, 56, and 57 incorporates the elements of independent claim 17.

Csicsatka does not make-up for the failure of Jun in view of Kulas to disclose and suggest the elements of independent claim 17 discussed above. Indeed, Csicsatka does not expressly or inherently disclose concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence. To the contrary, in accordance with Csicsatka's teachings the audio clips of an album are rendered without any type of transition between the audio clips.

Therefore, claims 50, 51, 54, 56, and 57 are patentable over Jun in view of Kulas and Csicsatka for at least the same reasons explained above in connection with independent claim 17.

3. Independent claim 15

Independent claim 15 recites:

15. An audio processing method, comprising:  
sequentially rendering audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each audio summary comprises digital content summarizing at least a portion of a respective associated audio piece, wherein each audio

piece is associated with multiple audio summaries and a single audio summary is rendered automatically for each audio piece; and

rendering an audio summary for a given audio piece in response to user input received during rendering of a preceding audio summary associated with the given audio piece.

The rejection of independent claim 15 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Csicsatka should be withdrawn because the cited references, taken alone or in any permissible combination, do not disclose or suggest each and every one of the elements of the claim.

The Examiner has acknowledged that neither Jun nor Kulas discloses or suggests that "each audio piece is associated with multiple audio summaries and a single audio summary is rendered automatically for each audio piece" (see page 12, last full ¶). The Examiner has relied on Csicsatka to make-up for the failure of Jun in view of Kulas to disclose this element of claim 15. In particular, the Examiner has stated that (pages 13-14 of the final Office action; emphasis added):

Regarding a single audio summary is rendered automatically for each audio piece; and rendering an audio summary for a given audio piece in response to user input received during rendering of a preceding audio summary associated with the given audio piece; Csicsatka discloses a similar audio processing method (Method of Creating Playlists using Audio Clips), comprising: identifying audio summaries ("Audio Clips") of respective audio pieces ("songs"), wherein each of the audio summaries comprises digital content summarizing at least a portion of the respective audio piece ("Audio Clips can be used to identify content without having to play the entire selection or look at a display," page 1 sixth paragraph). Csicsatka further discloses, "The user selects an Album. Audio Clips from the songs from the album start to play in album order. While the Audio Clip is playing a press of the Favorites Key will add it to the playlist. Once it has been added the player advances to the next song. If the Audio Clip is set to 5 seconds and there are 12 songs on the album a playlist could be created in less than 60 seconds," page 1 example 1.

On its face, the Examiner's statement does not show that Csicsatka discloses "each audio piece is associated with multiple audio summaries and a single audio summary is rendered

automatically for each audio piece.” Instead, the Examiner’s statement merely shows that audio clips start to play in album order. Thus, the Examiner has not established a *prima facie* case that claim 15 is obvious.

Moreover, Csicsatka does not disclose or suggest that “each audio piece is associated with multiple audio summaries,” as recited in claim 15. Instead, Csicsatka discloses that each song is represented by a respective audio clip that has a start time and a duration (see, e.g., Example 1 on page 1 and the description under the “Audio Clips” bullet item on page 2 of the provisional application). Therefore, none of the cited references discloses or suggests that “each audio piece is associated with multiple audio summaries and a single audio summary is rendered automatically for each audio piece,” as recited in claim 1.

For at least these reasons, the rejection of independent claim 15 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Csicsatka should be withdrawn.

The rejection of claim 15 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Herley also should be withdrawn because at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas and Csicsatka in the manner proposed by the Examiner.

The Examiner has taken the position that (see § 6 on pages 16-17 of the final Office action):

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to apply the transition audio segments taught by Kulas in the digest streams taught by Jun, since the transitional audio segments are used "to indicate the end of a sample during playback mode and to indicate the start of a next sample," Kulas Column 5 Lines 52 -65, thereby allowing the user do distinguish when one sample ends and the other begins.

As explained above, Jun discloses a method in which digest segments of a multimedia stream of a particular multimedia content (e.g., movie, drama, sports; see col. 1, line 29) are identified and ranked into different levels of a digest segment hierarchy (see col. 4, lines 20-29, and FIG. 1), where each digest segment is a sub-stream of the original multimedia stream (see col. 1, lines 35-36 and 56-57). Kulas discloses outputting a tone “to indicate the end of a sample during sample playback mode and to indicate the start of a next sample” (col. 5, lines 53-56).

Thus, the Examiner's proposed combination of Jun and Kulas would result in a audio tone being played at the end of each digest segment extracted from a particular multimedia content.

For the reasons explained below, one skilled in the art the time the invention was made would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner.

First, one skilled in the art would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner because the proposed combination would not have served any readily apparent purpose. Jun discloses that digest segments of an audio-visual multimedia stream (e.g., movie, drama, sports; see col. 1, line 29) are identified and ranked into different levels of a digest segment hierarchy (see col. 4, lines 20-29, and FIG. 1), where each digest segment is a sub-stream of the original multimedia stream (see col. 1, lines 35-36 and 56-57). The audio-video multimedia streams disclosed in Jun are quite difference from the audio pieces disclosed in Kulas. In particular, in the case of audio pieces, it is not readily readily apparent to the listener when a four-second audio sample of an audio piece has ended and a four-second audio sample of a successive audio piece has begun because the audio samples of different audio pieces from the same CD oftentimes sound very similar to one another. This is the reason Kulas inserts a tone between the audio samples. In the case of audio-video content, on the other hand, the video content itself provides all the cues needed to inform the viewer when a digest segment has ended because the video contents of successive digest segments selected from the same multimedia stream are quite likely to be visually different. This is the reason Jun does not disclose or suggest anything whatsoever about inserting audible tones between successive digest segments.

Second, one skilled in the art would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner because the proposed combination would produce an undesirable result. In particular, playing a tone as taught by Kulas between the audio-visual digest segments of a particular multimedia stream as taught by Jun would be quite discordant to the viewer who is browsing through sequences of the digest segments because such audio interruptions are not necessary and therefore are not expected in the context of audio-video browsing.

For at least this additional reason, the rejection of independent claim 15 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Csicsatka should be withdrawn.

4. Independent claim 54

Independent claim 54 recites:

54. An audio processing system, comprising:

a rendering engine operable to sequentially render audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each audio piece is associated with multiple audio summaries, the rendering engine is operable to render a single audio summary automatically for each audio piece, and the rendering engine additionally is operable to render an audio summary for a given audio piece in response to user input received during rendering of a preceding audio summary associated with the given audio piece..

Independent system claim 54 recites elements that essentially track the pertinent features of independent claim 15 discussed above. Therefore, claim 54 is patentable over Jun in view of Kulas and Csicsatka for at least the same reasons explained above in connection with independent claim 15. In particular, the rejection of independent claim 54 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Csicsatka should be withdrawn because Jun, Kulas, and Csicsatka, taken either alone or in any permissible combination, do not disclose or suggest all the elements of the claimed invention. For example, for the reasons explained above in connection with independent claim 15, neither Jun nor Kulas nor Csicsatka discloses or suggests “wherein each audio piece is associated with multiple audio summaries, [and] the rendering engine is operable to render a single audio summary automatically for each audio piece,” as recited in claim 54. The rejection of independent claim 54 also should be withdrawn because at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas and Csicsatka in the manner proposed by the Examiner for the reasons explained above in connection with independent claim 15.



**D. The rejection of claims 13, 14, 52, and 53 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Herley**

The Examiner has rejected claims 13, 14, 52, and 53 under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Herley (U.S. 2004/0002310).

**1. Independent claim 13**

Independent claim 13 recites:

13. An audio processing method, comprising:
- sequentially rendering audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each of the audio summaries comprises digital content summarizing at least a portion of a respective associated audio piece;
  - receiving a user request to browse the audio summaries;
  - ordering ones of the audio summaries into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was received; and
  - rendering the sequence.

The rejection of independent claim 13 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Herley should be withdrawn because the cited references, taken alone or in any permissible combination, do not disclose or suggest each and every one of the elements of the claim.

The Examiner has acknowledged that neither Jun nor Kulas discloses or suggests “ordering ones of the audio summaries into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was received,” as recited in claim 13 (see § on pages 15-18 of the final Office action). The Examiner has relied on Herley in an effort to make-up for the failure of Jun in view of Kulas to disclose or suggest the “ordering” element of claim 13 (see § 6 on pages 17-18 of the final Office action). Contrary to the Examiner’s position, however, Herley does not disclose or suggest “ordering ones of the audio summaries into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was received.”

In particular, the Examiner has taken the position that Herley discloses the "ordering" element of claim 13 in ¶¶ 68, 70, 86, 91, 92, and 116. The cited sections of Herley's disclosure, however, teach that the similarities between media items may be used to generate a sub-collection for use by the scan list engine (see ¶ 86) and that scan list engine play segments of the media items (or seed items) in the sub-collection as follows: a first media item of the sub-collection is randomly selected; a second element of the sub-collection is chosen so as to be substantially dissimilar to the first element; and subsequent elements are iteratively chosen so as to be dissimilar to several previous elements (see ¶ 50). Thus, in the context of the scan entertainment content function relied on by the Examiner in support of the rejection of claim 13, Herley expressly discloses playing the media item segments in order of dissimilarity.

Thus, contrary to the Examiner's position Herley expressly teaches away from "ordering ones of the audio summaries into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was received," as recited in claim 13. Therefore, none of the cited references discloses or suggests the "ordering" element of claim 13.

For at least these reasons, the rejection of independent claim 13 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Herley should be withdrawn.

The rejection of claim 13 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Herley also should be withdrawn because at the time the invention was made there was no apparent reason to combine the teachings of Jun in view of Kulas and Herley in the manner proposed by the Examiner.

The Examiner has taken the position that (see § 6 on pages 16-17 of the final Office action):

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to apply the transition audio segments taught by Kulas in the digest streams taught by Jun, since the transitional audio segments are used "to indicate the end of a sample during playback mode and to indicate the start of a next sample," Kulas Column 5 Lines 52 -65, thereby allowing the user do distinguish when one sample ends and the other begins.

As explained above, Jun discloses a method in which digest segments of a multimedia stream of a particular multimedia content (e.g., movie, drama, sports; see col. 1, line 29) are identified and ranked into different levels of a digest segment hierarchy (see col. 4, lines 20-29, and FIG. 1), where each digest segment is a sub-stream of the original multimedia stream (see col. 1, lines 35-36 and 56-57). Kulas discloses outputting a tone "to indicate the end of a sample during sample playback mode and to indicate the start of a next sample" (col. 5, lines 53-56). Thus, the Examiner's proposed combination of Jun and Kulas would result in a audio tone being played at the end of each digest segment extracted from a particular multimedia content.

For the reasons explained below, one skilled in the art the time the invention was made would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner.

First, one skilled in the art would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner because the proposed combination would not have served any readily apparent purpose. Jun discloses that digest segments of an audio-visual multimedia stream (e.g., movie, drama, sports; see col. 1, line 29) are identified and ranked into different levels of a digest segment hierarchy (see col. 4, lines 20-29, and FIG. 1), where each digest segment is a sub-stream of the original multimedia stream (see col. 1, lines 35-36 and 56-57). The audio-video multimedia streams disclosed in Jun are quite difference from the audio pieces disclosed in Kulas. In particular, in the case of audio pieces, it is not readily readily apparent to the listener when a four-second audio sample of an audio piece has ended and a four-second audio sample of a successive audio piece has begun because the audio samples of different audio pieces from the same CD oftentimes sound very similar to one another. This is the reason Kulas inserts a tone between the audio samples. In the case of audio-video content, on the other hand, the video content itself provides all the cues needed to inform the viewer when a digest segment has ended because the video contents of successive digest segments selected from the same multimedia stream are quite likely to be visually different. This is the reason Jun does not disclose or suggest anything whatsoever about inserting audible tones between successive digest segments.

Second, one skilled in the art would not have had any apparent reason to combine the cited references in the manner proposed by the Examiner because the proposed combination

would produce an undesirable result. In particular, playing a tone as taught by Kulas between the audio-visual digest segments of a particular multimedia stream as taught by Jun would be quite discordant to the viewer who is browsing through sequences of the digest segments because such audio interruptions are not necessary and therefore are not expected in the context of audio-video browsing.

For at least this additional reason, the rejection of independent claim 13 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Herley should be withdrawn.

2. Dependent claim 14

Claim 14 incorporates the elements of independent claim 13 and therefore is patentable over Jun in view of Kulas and Herley for at least the same reasons explained above.

3. Independent claim 52

Independent system claim 52 recites elements that essentially track the pertinent features of independent claim 13 discussed above. Therefore, claim 52 is patentable over Jun in view of Kulas and Herley for at least the same reasons explained above in connection with independent claim 13.

4. Dependent claim 53

Claim 53 incorporates the elements of independent claim 52 and therefore is patentable over Jun in view of Kulas and Herley for at least the same reasons explained above.

**E. The rejection of claim 16 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Takenaka**

The Examiner has rejected claim 16 under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Takenaka (U.S. 6,807,450).

Claim 16 incorporates elements of independent claim 1. Takenaka does not make-up for the failure of Jun in view of Kulas to disclose or suggest the elements of independent claim 1 discussed above. Therefore, claim 16 is patentable over Jun in view of Kulas and Takenaka for at least the same reasons explained above in connection with independent claim 1.

Claim 16 also is patentable over Jun in view of Kulas for the following additional reason.

Claim 16 recites "normalizing audio summaries to a common loudness level."

The Examiner has acknowledged that neither Jun nor Kulas discloses or suggests the elements of claim 16. Instead, the Examiner has relied on the teachings of Takenaka as follows (see page 19 of the final Office action):

...Takenaka discloses a digital audio reproduction method (Fig. 5E; column 12, lines 11 -29) that provides transition audio segments between the information pieces (i.e., music items). Takenaka further discloses that such an arrangement provides a natural linkage between songs, enhancing listener enjoyment (column 12, lines 42 -45) and further discloses reproduction at a constant level (i.e., normalizing to a common loudness level) (Fig 5E; column 12, lines 30 -36). ...

Contrary to the Examiner's statement, however, Takenaka does not teach or suggest anything whatsoever about normalizing audio data to a common loudness level. In FIG. 5E, Takenaka merely shows the attenuation levels that are applied by the DSP 14 to avoid "unnatural" transitions during the playback of consecutive songs, the second embodiment (see, e.g., the "Amount of Attenuation" label in FIG. 5E, and col. 11, lines 39-62). In column 12, lines 11-49, Takenaka merely discloses the amount by which the DSP 14 attenuates the digital audio data; Takenaka does not disclose anything about normalizing the digital audio data to a common loudness level.

For this additional reason, the Examiner's rejection of claim 16 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Takenaka should be withdrawn.

**F. The rejection of claim 49 under 35 U.S.C. § 103(a) over Jun in view of Kulas and Setogawa**

The Examiner has rejected claim 49 under 35 U.S.C. § 103(a) over Jun (U.S. 6,918,081) in view of Kulas (U.S. 6,044,047) and Setogawa (U.S. 6,424,793).

Claim 49 incorporates elements of independent claim 1. Setogawa does not make-up for the failure of Jun in view of Kulas to disclose or suggest the elements of independent claim 1 discussed above.

Applicant : Tong Zhang et al.  
Serial No. : 10/611,449  
Filed : July 1, 2003  
Page : 30 of 40

Attorney's Docket No.: 100202720-1  
Appeal Brief dated Oct. 8, 2008  
Reply to final Office action dated July 8, 2008

Therefore, claim 49 is patentable over Jun in view of Kulas and Setogawa for at least the same reasons explained above in connection with independent claim 1.

#### VIII. Conclusion

For the reasons explained above, all of the pending claims are now in condition for allowance and should be allowed.

Charge any excess fees or apply any credits to Deposit Account No. 08-2025.

Respectfully submitted,

Date: Oct. 8, 2008

/Edouard Garcia, Reg. No. 38,461/

Edouard Garcia

Reg. No. 38,461

Telephone No.: (650) 965-8342

Please direct all correspondence to:

Hewlett-Packard Company  
Intellectual Property Administration  
Legal Department, M/S 35  
P.O. Box 272400  
Fort Collins, CO 80528-9599

### CLAIMS APPENDIX

The following Listing of Claims replaces all prior versions, and listings, of claims in the application.

#### Listing of Claims:

Claim 1 (previously presented): An audio processing method, comprising:  
identifying audio summaries of respective audio pieces, wherein each of the audio summaries comprises digital content summarizing at least a portion of the respective audio piece, and the identifying comprises for each of the audio pieces  
selecting constituent segments of the audio piece as its respective ones of the audio summaries and  
ranking its audio summaries into different levels of a respective audio summary hierarchy;  
determining transition audio segments each comprising a form of audio content that is different from the audio summaries and distinguishes the transition audio segment from the audio summaries;  
concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence in which at least one of the transition audio segments is between successive ones of the audio summaries; and  
rendering the sequence.

Claim 2 (original): The method of claim 1, wherein identical transition audio segments are rendered between pairs of sequential audio summaries.

Claim 3 (previously presented): An audio processing method, comprising:  
sequentially rendering audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each audio summary comprises digital content summarizing at least a portion of a respective

associated audio piece, wherein identical transition audio segments are rendered between pairs of sequential audio summaries and each identical transition audio segment corresponds to a Gabor function in a time domain representation.

Claim 4 (previously presented): The method of claim 1, wherein each of the transition audio segments corresponds to a Gabor function in a time domain representation, and each of the Gabor functions has a respective center frequency substantially corresponding to a center pitch of an adjacent one of the audio summaries in the sequence.

Claim 5 (original): The method of claim 1, wherein the audio summaries and the interleaved transition audio segments are rendered consecutively.

Claim 6 (original): The method of claim 1, wherein each audio summary is a representative segment of a respective associated audio piece.

Claim 7 (original): The method of claim 1, further comprising classifying audio pieces into categories in response to user input received during rendering of the associated audio summaries.

Claim 8 (original): The method of claim 7, further comprising building a playlist based on categories assigned to a set of audio pieces.

Claim 9 (previously presented): The method of claim 1, wherein at least one audio summary is linked to an associated audio piece by a browsable link.

Claim 10 (previously presented): The method of claim 9, further comprising rendering a given one of the audio pieces linked by a browsable link to an associated one of the audio summaries in response to user input received during rendering of the associated audio summary, wherein the rendering comprises following the browsable link from the associated audio



summary to the given audio piece before rendering a successive one of the transition audio segments.

Claim 11 (previously presented): The method of claim 1, further comprising rendering a given audio piece beginning at a location in the given audio piece linked by a browsable link to an audio summary associated with the given audio piece, wherein the rendering comprises following the browsable link from the associated audio summary to the given audio piece.

Claim 12 (original): The method of claim 11, further comprising rendering a second audio piece at a location in the second audio piece linked to a successive audio summary associated with the second audio piece.

Claim 13 (previously presented): An audio processing method, comprising:  
sequentially rendering audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each of the audio summaries comprises digital content summarizing at least a portion of a respective associated audio piece;  
receiving a user request to browse the audio summaries;  
ordering ones of the audio summaries into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was received; and  
rendering the sequence.

Claim 14 (original): The method of claim 13, wherein audio summaries are rendered in accordance with the ordered sequence.

Claim 15 (previously presented): An audio processing method, comprising:  
sequentially rendering audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each audio summary comprises digital content summarizing at least a portion of a respective

associated audio piece, wherein each audio piece is associated with multiple audio summaries and a single audio summary is rendered automatically for each audio piece; and

rendering an audio summary for a given audio piece in response to user input received during rendering of a preceding audio summary associated with the given audio piece.

Claim 16 (original): The method of claim 1, further comprising normalizing audio summaries to a common loudness level.

Claim 17 (previously presented): An audio processing system, comprising a rendering engine operable to perform operations comprising:

identifying audio summaries of respective audio pieces, wherein each of the audio pieces is associated with respective ones of audio summaries ranked into different levels of a respective audio summary hierarchy and in the identifying the rendering engine is operable to identify the respective levels into which the audio summaries are ranked;

determining transition audio segments each comprising a form of audio content that is different from the audio summaries and distinguishes the transition audio segment from the audio summaries;

concatenating the transition audio segments and ones of the audio summaries ranked at a selected level of the audio summary hierarchies into a sequence in which at least one of the transition audio segments is between sequential ones of the audio summaries; and  
rendering the sequence.

Claims 18-38 (canceled)

Claim 39 (previously presented): The method of claim 1, further comprising following links between multiple ones of the audio summaries and one of the audio pieces.

Claim 40 (previously presented): The method of claim 1, wherein each of the transition audio segments corresponds to a monotone sound.

Claim 41 (previously presented): The method of claim 1, wherein the rendering comprises rendering the audio summaries and the transition audio segments consecutively without any gaps between the audio summaries and the transition audio segments.

Claim 42 (previously presented): The method of claim 1, further comprising, in response to user input during rendering of a current one of the audio summaries that comprises digital content summarizing at least a portion of a given one of the audio pieces, rendering another audio summary in a hierarchical cluster of audio summaries each of which comprises digital content summarizing at least a portion of the given audio piece, wherein the hierarchical cluster includes the current audio summary.

Claim 43 (previously presented): The method of claim 1, further comprising receiving one or more user-specified categories for respective ones of the audio summaries while the audio summaries and the transition audio segments are being rendered.

Claim 44 (previously presented): The method of claim 43, further comprising building one or more playlists based on the one or more user-specified categories.

Claim 45 (previously presented): The method of claim 1, wherein at least one of the audio summaries is associated with a pointer to a location in a respective one of the audio pieces.

Claim 46 (previously presented): The method of claim 1, further comprising following a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the specified location.

Claim 47 (previously presented): The method of claim 46, further comprising terminating the rendering of the associated audio piece and resuming the sequential rendering of the audio summaries and the transition audio segments.

Claim 48 (previously presented): The method of claim 47, wherein the terminating is initiated in response to user input.

Claim 49 (previously presented): The method of claim 47, wherein the terminating is initiated in response to completion of the rendering of the associated audio piece.

Claim 50 (previously presented): The system of claim 17, wherein the rendering engine is operable to assign user-specified categories to respective ones of the audio pieces in response to user input.

Claim 51 (previously presented): The system of claim 50, wherein the rendering engine is operable to build a playlist based on the user-specified categories assigned to the ones of the audio pieces.

Claim 52 (previously presented): An audio processing system, comprising:  
a rendering engine operable to sequentially render audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each of the audio summaries comprises digital content summarizing at least a portion of a respective associated audio piece and, in response to receipt of a user request to browse the audio summaries, the rendering engine is operable to order ones of the audio summaries into a sequence in order of audio feature vector closeness to a given one of the audio summaries being rendered when the user request was received and to render the sequence.

Claim 53 (previously presented): The system of claim 52, wherein the rendering engine is operable to render the audio summaries in accordance with the ordered sequence.

Claim 54 (previously presented): An audio processing system, comprising:  
a rendering engine operable to sequentially render audio summaries and transition audio segments with at least one transition audio segment rendered between each pair of sequential audio summaries, wherein each audio piece is associated with multiple audio summaries, the

rendering engine is operable to render a single audio summary automatically for each audio piece, and the rendering engine additionally is operable to render an audio summary for a given audio piece in response to user input received during rendering of a preceding audio summary associated with the given audio piece.

Claim 55 (previously presented): The system of claim 17, wherein the rendering engine is operable to render the audio summaries and the transition audio segments consecutively without any gaps between the audio summaries and the transition audio segments.

Claim 56 (previously presented): The system of claim 17, wherein the rendering engine is operable to receive one or more user-specified categories for ones of the audio summaries while the audio summaries and the transition audio segments are being rendered.

Claim 57 (previously presented): The system of claim 56, wherein the rendering engine is operable to build one or more playlists based on the one or more user-specified categories.

Claim 58 (previously presented): The system of claim 17, wherein the rendering engine is operable to follow a pointer from a given audio summary being rendered to a location in an associated audio piece specified by the pointer, and rendering the associated audio piece beginning at the specified location.

Claim 59 (previously presented): The system of claim 58, wherein the rendering engine is operable to terminate the rendering of the associated audio piece and resume the sequential rendering of the audio summaries and the transition audio segments.

Claim 60 (previously presented): The system of claim 17, wherein each transition audio segment corresponds to a monotone sound.

Applicant : Tong Zhang et al.  
Serial No. : 10/611,449  
Filed : July 1, 2003  
Page : 38 of 40

Attorney's Docket No.: 100202720-1  
Appeal Brief dated Oct. 8, 2008  
Reply to final Office action dated July 8, 2008

Claim 61 (previously presented): The method of claim 1, wherein the rendering comprises rendering only one transition audio segment between each sequential pair of the audio summaries.

Claim 62 (previously presented). The method of claim 1, wherein the identifying comprises identifying the audio summaries of the audio pieces based on links between the audio pieces and respective ones of the audio summaries.

Applicant : Tong Zhang et al.  
Serial No. : 10/611,449  
Filed : July 1, 2003  
Page : 39 of 40

Attorney's Docket No.: 100202720-1  
Appeal Brief dated Oct. 8, 2008  
Reply to final Office action dated July 8, 2008

### EVIDENCE APPENDIX

There is no evidence submitted pursuant to 37 CFR §§ 1.130, 1.131, or 1.132 or any other evidence entered by the Examiner and relied upon by Appellant in the pending appeal. Therefore, no copies are required under 37 CFR § 41.37(c)(1)(ix) in the pending appeal.

Applicant : Tong Zhang et al.  
Serial No. : 10/611,449  
Filed : July 1, 2003  
Page : 40 of 40

Attorney's Docket No.: 100202720-1  
Appeal Brief dated Oct. 8, 2008  
Reply to final Office action dated July 8, 2008

RELATED PROCEEDINGS APPENDIX

Appellant is not aware of any decisions rendered by a court or the Board that will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal. Therefore, no copies are required under 37 CFR § 41.37(c)(1)(x) in the pending appeal.